2013

# High-throughput phenotyping of uropathogenic E. coli isolates with Fourier transform infrared spectroscopy

AlRabiah, H



10.1039/c3an36517d
The Analyst
Royal Society of Chemistry (RSC)

Analyst

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

# High-throughput phenotyping of uropathogenic *E. coli* isolates with Fourier transform infrared spectroscopy†

Cite this: *Analyst*, 2013, **138**, 1363

Haitham AlRabiah,[a] Elon Correa,[a] Mathew Upton[b] and Royston Goodacre*[a]

Fourier transform infrared (FT-IR) spectroscopy is an established rapid whole-organism fingerprinting method that generates metabolic fingerprints from bacteria that reflect the phenotype of the microorganism under investigation. However, whilst FT-IR spectroscopy is fast (typically 10 s to 1 min per sample), the approaches for microbial sample preparation can be time consuming as plate culture or shake flasks are used for growth of the organism. We report a new approach that allows micro-cultivation of bacteria from low volumes (typically 200 μL) to be coupled with FT-IR spectroscopy. This approach is fast and easy to perform and gives equivalent data to the lengthier and more expensive shake flask cultivations (sample volume = 20 mL). With this micro-culture approach we also demonstrate high reproducibility of the metabolic fingerprints. The approach allowed separation of different isolates of *Escherichia coli* involved in urinary tract infection, including members of the globally disseminated ST131 clone, with respect to both genotype and resistance or otherwise to the antibiotic Ciprofloxacin.

Received 18th October 2012
Accepted 11th January 2013

DOI: 10.1039/c3an36517d

www.rsc.org/analyst



## Introduction

One of the most common infections in humans, and in females in particular, is urinary tract infection (UTI). Research has led to the assumption that worldwide UTI will affect approximately 50% of women at some point in their lives.[1,2] Of these approximately 50% will have recurrent UTI and some of these women will suffer from chronic UTIs.[3] *Escherichia coli* is the microorganism most commonly contributing to UTI (both community and nosocomial (an infection acquired whilst in hospital)).[4] In an international survey of midstream urine samples taken at 252 centres in 17 countries it was found that 77% of all isolates, 80% of general infections and 40% of nosocomial infections could be attributed to *E. coli*.[5,6] More recent molecular epidemiological analyses indicate that a relatively limited number of clones of uropathogenic *E. coli* (UPEC) cause the majority of UTI and members of these clones are often multi-drug resistant.[7] A particularly important clone is sequence type (ST)131, which is globally disseminated.[8]

In order to identify UTI pathogens traditional biochemical methods based on growth and nutritional properties are used. However, in general these are protracted (based on the growth rate of the pathogen), which can compromise the effectiveness of treatment of infections. Whilst recent molecular techniques have led to more rapid solutions for characterising microorganisms, unfortunately they tend to rely on already know DNA sequences for identification and the costs incurred by such methods, coupled with the highly specialised equipment render them impractical for use in routine laboratories.[9,10] Therefore fast, accurate, automated methods that are relatively inexpensive need to be developed. The ideal technique would be successful in identifying sources of infection effectively as well as making differentiation to sub-species level possible, for epidemiological purposes. One possible solution for this so-called whole organism fingerprinting is Fourier transform infrared (FT-IR) spectroscopy, which is an effective technique for bacterial identification and discrimination. The advantages of this method are that it is non-destructive and can create distinct spectral fingerprints for different microorganisms.[11–13] The FT-IR spectra generated from bacteria can be used to examine cell components such as proteins, nucleic acids, carbohydrates and lipids, and are manifest due to the vibrational modes of different functional group complements within these biochemical classes.[14] Research has shown that FT-IR spectroscopy, alongside multivariate data analysis, can give quick, user friendly, relatively inexpensive screening with the ability to differentiate bacteria at different taxonomic levels.[15–18]

Current methods for FT-IR require that the bacteria are grown and then sampled and this can be labour intensive when many different bacterial isolates need to be analysed. Whilst Naumann and colleagues have increased the speed of analysis using micro-cultivation and infrared analysis *via* FT-IR microscopy, this unfortunately requires expensive instrumentation and quite detailed image analysis for automation.[19] Thus, the aim of this

[a]*School of Chemistry and Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester, M1 7DN, UK. E-mail: roy.goodacre@manchester.ac.uk*

[b]*School of Medicine, University of Manchester, Stopford Building, Oxford Road, Manchester, M13 9PL, UK*

† Electronic supplementary information (ESI) available: See DOI: 10.1039/c3an36517d

study was to develop a high throughput methodology that readily couples bacterial growth requirements with FT-IR spectroscopy. This approach we illustrate for the discrimination between UPEC isolates with different genetic sequence types and different susceptibilities to quinolones; a group of antibiotics that inhibit nucleic acid synthesis directly by targeting DNA gyrase and topoisomerase IV. The developed method is rapid, less labour intensive than larger scale approaches and up to 200 samples can be simultaneously analysed under identical growth conditions, which is a prerequisite for whole organism fingerprinting methods that measure bacterial phenotype rather than genotype *per se*.

## Materials and methods

### 1.1 Microorganisms

Ten UPEC isolates were recovered from urine samples submitted to the bacteriology laboratory at Central Manchester Foundation Trust. These consisted of two isolates that were quinolone sensitive (isolates 173 and 191 of ST131) and eight isolates that were quinolone resistant (isolates 160, 163 and 164 of ST131; isolates 152, 161, 162, 169 and 171 that were non-ST131).

Each isolate was streaked on an agar plate to obtain axenic colonies. Biomass was collected from these single colonies to prepare 1 mL 20% glycerol working inocula stocks, which were stored at −20 °C. Isolates were routinely cultured in LB medium at 37 °C for 18 h.

### 1.2 Media

Lysogeny broth (LB) was prepared by dissolving 10 g of tryptone (Formedia, Hunstanton, UK), 5 g of yeast extract (USP, Cleveland, USA) and 10 g of sodium chloride (Fisher Scientific Ltd, Loughborough, UK), in 1 L of reverse osmosis water and after this it was autoclaved (121 °C, 15 min and 15 psi).

### 1.3 Antibiotics

In order to calculate the minimum inhibitory concentration (MIC) for Ciprofloxacin, 100 mg of the hydrochloride salt (Discovery fine chemicals, Dorset, UK) was dissolved in 50 mL distilled water. The final concentrations tested were: 100, 80, 50, 40, 25, 20, 12.5, 10, 8, 6.25, 5, 3.125, 2.5, 1.56, 1.25, 1, 0.78, 0.625, 0.5, 0.39, 0.3125, 0.3, 0.25, 0.1563, 0.1, 0.078, 0.05, 0.039, 0.03, 0.025, 0.02, 0.0195, and 0.0025 mg $L^{-1}$ according to ref. 20.

### 1.4 Growth conditions

For the start of each experiment, 49 mL of medium was inoculated with 1 mL of working stock from the freezer and incubated at 37 °C, in a shaking incubator at 200 rpm for 24 h. After incubation overnight, cultures (1 mL) were diluted with 49 mL fresh media and further incubated at 37 °C, 200 rpm for 1 h. These new axenic cultures were diluted to 0.5 McFarland standard at optical density (OD) 600 nm using Biomate 5 (Thermo, Hemel Hempstead, UK) and used as experimental inocula.

In order to standardise the inocula for growth curve experiments, 10 mL of the culture was diluted and washed two times with physiological saline (0.9% NaCl). The bacterial turbidity was adjusted to be equivalent to a 0.5 McFarland standard (OD 0.1 ± 0.02) at optical density (OD) 600 nm using a Biomate 5 (Thermo, Hemel Hempstead, UK).

The bacterial growth curves (typically $n = 5$ per experiment) were measured at OD 600 nm in a bioscreen spectrophotometer (Labsystems, Basingstoke, UK). This 'bioscreen' was run at the following settings: 10 min preheating, then an incubation temperature of 37 °C, with continuous medium shake, measurement interval of 10 min and in general 18 h for the total experiment. Typical growth curves, including when bacteria were exposed to antibiotics, are shown in Fig. S1.†

### 1.5 FT-IR spectroscopy

**1.5.1 Sample preparation.** Five different methods of sample preparation were investigated in order to develop a high-throughput method for discrimination between *E. coli* isolates of different clonal lineages; see ESI Fig. S2 and S3† for cartoons of this process.

*1.5.1.1 Method 1.* 19 mL of LB medium was inoculated with 1 mL of 0.5 McFarland standard equivalent inocula of pathogenic isolates (see Section 1.4) in 100 mL flasks and incubated for 18 h at 37 °C and 200 rpm. 450 μL from each culture was collected and the biomass was washed 3 times with normal saline and re-suspended in 400 μL of normal saline. 20 μL was spotted onto a zinc selenide (ZnSe) plate (Bruker Ltd, Coventry, UK) and oven dried at 40 °C for 45 min.

*1.5.1.2 Method 2.* 19 mL of LB medium was inoculated with 1 mL of 0.5 McFarland standard inocula of pathogenic isolates in 100 mL flasks and incubated for 18 h at 37 °C and 200 rpm. 20 μL from each flask was spotted directly onto a ZnSe plate and oven dried at 40 °C for 45 min.

*1.5.1.3 Method 3.* 190 μL of LB medium was inoculated with 10 μL of 0.5 McFarland standard inocula of pathogenic isolates (see Section 1.4) in a bioscreen plate and incubated as described in Section 1.4. In order to produce enough biomass each condition was cultured in 5 wells and 150 μL from each well was collected and mixed. 20 μL from each well was collected and spotted directly on to a ZnSe plate and oven dried at 40 °C for 45 min.

*1.5.1.4 Method 4.* This was identical to Method 3 except that after collection the biomass was washed 3 times with normal saline and re-suspended in 400 μL of normal saline. 20 μL was spotted onto a ZnSe plate and oven dried at 40 °C for 45 min.

*1.5.1.5 Method 5.* The supernatants from Method 4 were also collected and 20 μL was spotted on a ZnSe plate and oven dried at 40 °C for 45 min.

**1.5.2 Instrumentation.** Prior to analysis 96-well ZnSe plates were washed with 10% sodium dodecyl sulfate (SDS) solution, rinsed three times with deionised water and then rinsed three times with analytical grade propan-2-ol. The plates were finally rinsed with deionised water and air-dried at room temperature.

High throughput screening (HTS) FT-IR spectroscopic analysis was carried out using a Bruker Equinox 55 infrared spectrometer (Bruker Ltd, Coventry, UK) equipped with an HTX™ module using previously published methods.[21] FT-IR spectra were recorded directly from the dried cell biomass in transmission mode using a deuterated triglycine sulfate (DTGS) detector. A background spectrum was collected for each measurement from the reference well position (A1, 94 104 μm, 18 720 μm) of the ZnSe plates. All spectra were obtained in the

 

4000–600 $cm^{-1}$ range, and 64 scans were acquired at 4 $cm^{-1}$ resolution. These experimental conditions were maintained during all measurements. Spectra acquisition and spectral background subtractions were performed using OPUS software (Bruker Ltd, UK). The FTIR data were converted and analysed by MATLAB 2010a (The Mathworks Inc., Natwick, US) and R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria).

**1.5.3 Data analysis.** FT-IR data were first baseline corrected according to the extended multiplicative signal correction (EMSC) scaling method[22] then $CO_2$ signals from (i) 2400–2275 $cm^{-1}$ were removed from the spectra (and filled with a trend) and (ii) below 700 $cm^{-1}$ were also removed[21] because they might be considered to be misleading by causing spurious variations on multivariate analysis. The EMSC method was originally developed to reduce the disturbing effect of light-scattering, due to small particles scattering light more than larger ones.[23] This type of normalisation takes the information registered in the spectra and attempts to separate physical light-scattering effects from the actual light absorbed by molecules.

Data were then analysed using principal component analysis (PCA). PCA is an unsupervised method with no *a priori* knowledge of experimental structure and is used to reduce the dimensionality of the data. The objective of PCA is to explain the variance–covariance structure of a set of variables through a few linear combinations of these variables.[24] Much of the original data variability can be accounted for by a small number of principal components (PCs), which are then used for data reduction and visual data interpretation. The PCA results are discussed in terms of PCA scores and loadings; the PCs are the transformed variable values and the loadings are the weights by which the original data variables should be multiplied to obtain the component scores. After PCA, a supervised method, known as discriminant function analysis (DFA) was applied to the PCs (*i.e.* PC-DFA). PC-DFA depends on the prior knowledge of experimental structure (*i.e.*, the experimental class structure) and retained PCs to discriminate between groups (different classes).[25,26] DFA is a supervised technique that discriminates groups using *a priori* knowledge of class membership. The algorithm works to maximise between-group variance and minimise within-group variance.[27] In the present work, the PCs were used as inputs for DFA and the results were validated by 1000 bootstrap cross-validations. In this process 50 PC-DFA models were built which had from the 1$^{st}$ to the $n^{th}$ PC selected (where $n$ was set to a maximum of 50). Each of the models was rigorously tested using resampling methods to check that the clustering was not over-fitted. Bootstrap is a re-sampling technique that can be applied as cross-validation to estimate the prediction performance of a model. The basic idea of this method is to select randomly, with replacement, $N$ samples from a set containing exactly $N$ samples. All selected samples, including the repetitions, are then used as training set and the non-selected samples are used as test set.[28] One can think of this as having all samples analysed ($N = ||X||$ for our case) in a bag. A single sample is then taken out of the bag randomly and its number noted – this sample now forms part of the training data, and the sample is placed back into the bag. This random sample picking process is repeated until $||X||$ samples are in the training set. Some samples will be used multiple times, and on average 63.2% of all of the samples will have been selected for training, and the remaining 36.8% will be used as the test data.

For the best model, which was typically built using the first 30 PCs, 1000 bootstraps were conducted (*i.e.*, 1000 PC-DFA models were generated) and statistics performed on the test set only. We calculated 95% confidence interval (CI) and in the PC-DFA scores plots quote in parentheses the lower and upper bounds of the CI obtained for all 1000 models, we also provide the chi-square ($\chi^2$) statistics computed on the contingency table represented by the classification matrix, also known as confusion matrix, built over the 1000 models. The objective of this $\chi^2$ test was to compare the difference between the mean expected (real instances) and the mean observed (predicted instances) values in terms of the true positive, true negative, false positive and false negative cells of the classification matrix. The null hypothesis is that the expected and observed means are equal, which would suggest that the models are accurate. Therefore, the higher the $p$-value of the $\chi^2$ statistics (closer to 1) the stronger the evidence is to fail to reject the null hypothesis and thus the models are considered statistically valid.

Conversion of FT-IR absorbance spectra to an *XY* data matrix and the following multivariate data analyses were performed in MATLAB 2010a and R version 2.13. All scripts used for data analysis are available from the authors on request.

## Results and discussions

Pathogenic isolates of *E. coli* of different sequence types (ST131 and non-ST131) were challenged with different concentrations of Ciprofloxacin hydrochloride, which is completely soluble in water,[20] to determine the susceptibility to the drug and the MIC of each isolate (Fig. S1† displays example growth curves). It was found that isolates 160, 161, 163 and 169 were fully resistant, whilst isolates 173 and 191 were fully sensitive. In addition,

**Table 1** MIC of Ciprofloxacin hydrochloride for pathogenic isolates of *E. coli* of different sequence types (ST131 and non-ST131) with different susceptibilities to quinolones

| Isolate number | Sequence type | Quinolone phenotype | MIC range determination[a] (mg L) |
|---|---|---|---|
| 152 | non-ST131 | R | 0.25–0.1 |
| 161 | non-ST131 | R | No effect |
| 162 | non-ST131 | R | 12.5–10 |
| 169 | non-ST131 | R | No effect |
| 171 | non-ST131 | R | 0.5–0.25 |
| 160 | ST131 | R | No effect |
| 163 | ST131 | R | No effect |
| 164 | ST131 | R | 45–50 |
| 173 | ST131 | S | 0.03–0.02 |
| 191 | ST131 | S | 0.05–0.03 |

[a] The highest concentration used from Ciprofloxacin hydrochloride is 100 mg $L^{-1}$; R and S, indicate resistance or susceptibility to quinolones, respectively.

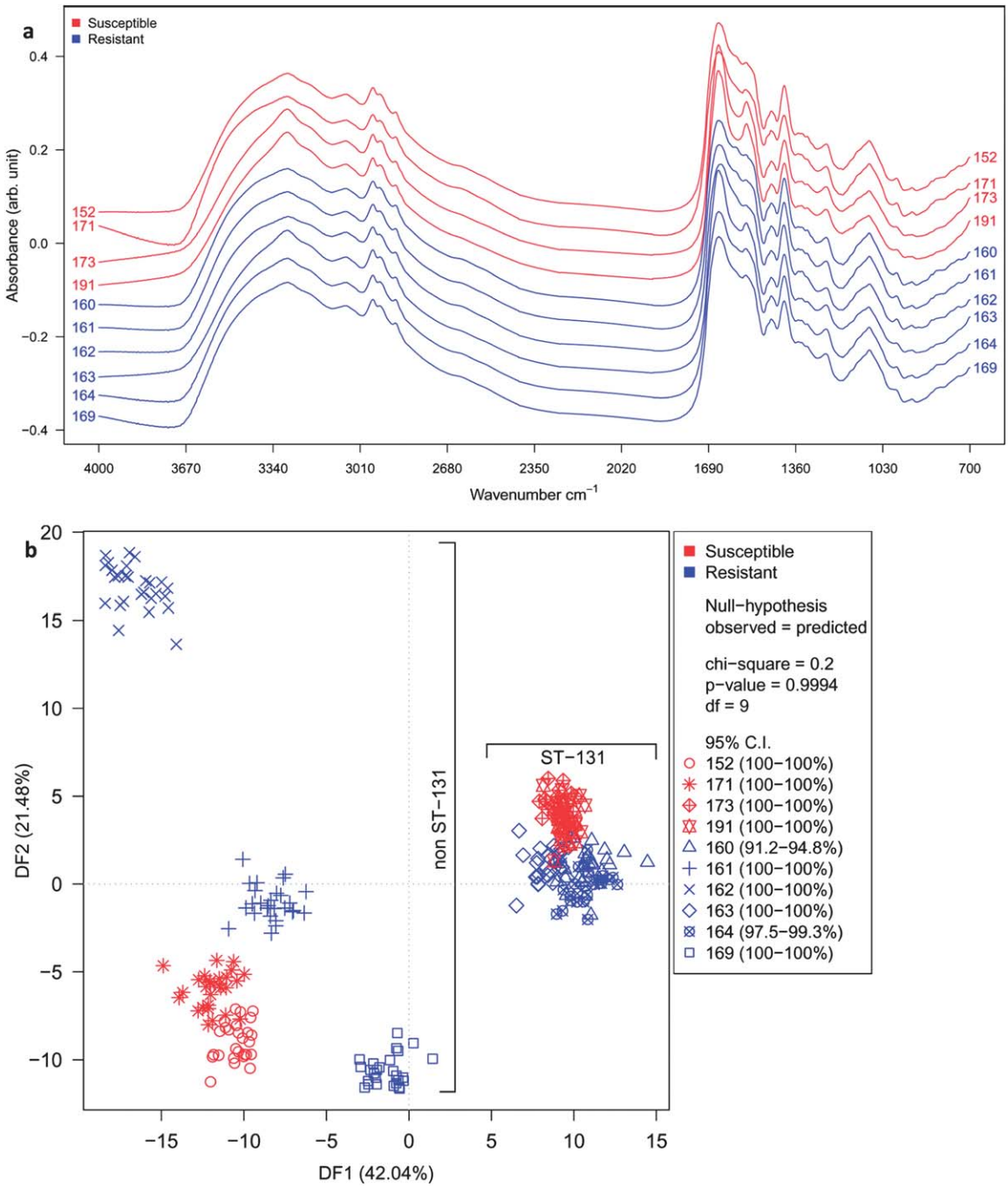


**Fig. 1** (a) Raw FT-IR spectra obtained from isolates 152, 160, 161, 162, 163, 164, 169, 171, 173 and 191 before washing the bacterial cells; the spectra are offset in the *Y*-axis for ease of visualisation. (b) PC-DFA scores plot of FT-IR spectra from the non-washed samples (30 PCs were extracted from PCA and used as inputs to DFA, these 30 PCs explain 99.97% of the totals explained variance); the legend beside the plot shows the 95% confidence interval for the 10 bacterial isolates estimated from the DFA model validation over 1000 independent bootstrap cross-validations.

isolates 152, 162, 164 and 171 showed different levels of susceptibility to Ciprofloxacin (Table 1).

Initial experiments in micro-culture on the bioscreen instrument were conducted using a modified method of sample preparation on just two isolates (162 and 163); these were chosen as they had different gene sequences and the results were compared with those of the standard shake flask method. The first step in modifying the conditions was the use of a bioscreen plate rather than flasks, which has some advantages over the usual method of preparing samples in shaking flasks. First, it is less laborious because the growth curves are generated directly and the same culture can then be used for analysis by FT-IR. This allows more confidence as growth profile and FT-IR spectra are from exactly the same biological replicate. It is also more accurate, because the automation process in the bioscreen eliminates errors and reduces the variation that might be produced by manual techniques. Another advantage is that the effort of the manual technique is reduced. Finally,

 

bioscreens have enough room to handle two 100-well sterile honeycomb plates, making it possible to run 200 samples at the same time under exactly the same conditions, which is difficult when following the classical microbial growth method.

To compare the new method and the shake flask method on isolates 162 and 163, each was cultured in multiple biological replicates both in bioscreen plates and in 100 mL flasks, under the same conditions. After growth the biomass was collected (Fig. S2 and S3†) and analysed using FT-IR spectroscopy either directly or after washing with physiological saline. The data from FT-IR were subjected to PCA and the four scores plots from each of the four different conditions are reported in the ESI (Fig. S4).† An additional plot of all cultures together (Fig. S5†) also helps illustrate the effects of washing and of culturing in bioscreen plates *versus* shake flasks. A clear distinction between these isolates regardless of the method of culturing can be observed and this is irrespective of the culture volume of the two methods: 20 mL for the flask cultures compared to just 200 μL in the bioscreen micro-culture instrument.

As the previous experiments showed that micro-culture and flasks gave equivalent clustering and considering the advantages of the micro-cultivation strategy this approach was explored further. Another series of experiments on ten isolates was conducted to determine the effects of biomass washing by analysing the samples in three different preparation methods: direct analysis (non-washed samples), washed samples, and the supernatants from the washed samples were also analysed by FT-IR spectroscopy and the resulting spectroscopic fingerprints compared using cluster analysis.

The FT-IR spectra from the direct analysis of the bacterial biomass is shown in Fig. 1a and these all display very similar features and this demonstrates the need for cluster analysis using PC-DFA. It can be see clearly from Fig. 1b that isolates of ST131 formed a definite and tight cluster that was located on the right (positive side) of DF1, while non-ST131 isolates were spread on the left (negative side) of DF1. This separation according to ST was encouraging from these non-washed samples and was not observed in either the supernatants (Fig. S6†) or washed samples (Fig. S7†). Further analysis of the STs individually from the non-washed samples (Fig. 2) revealed that there did appear to be some separation according to whether the *E. coli* isolates were resistant or otherwise to Ciprofloxacin. This was not observed in the washed bacterial cells or their supernatants (data not shown).

If we inspect Fig. 2 in more detail it can be readily observed that the sensitive isolates (red symbols) were generally closer to each other compared with the resistant isolates (blue symbols) in both ST131 and non-ST131 bacteria. With respect to ST131 isolates (Fig. 2a), another potential advantage of this FT-IR spectroscopic analysis is that the distribution of these isolates with the same genetic sequence might depend on their susceptibility to Ciprofloxacin since the sensitive ones are separated from resistant isolates in the second discriminant function (DF2).

In simple terms, the main differences between these sample preparation methods are that the direct analysis will measure everything: the bacteria themselves as well as any effect they have had on medium consumption and bacterial excretions (exometabolome). Whilst during the washing process these two factors are recovered and analysed separately. As the differences between the direct analysis and the other two sample preparation methods are so profound, they do require some explanation.

The first area where difference may occur is due changes to the medium during growth. The non-washed samples of isolates from both ST groups contained the media and the discrimination might be due to members of one ST preferentially consuming specific ingredients in the medium more than the other ST.[29] Whilst differentiation was not obviously manifest in the supernatant analysis (Fig. S6†), combination of the supernatant with the cells may have contributed to spectral differences clearly observed with respect to ST (Fig. 2). It has been reported previously that different *E. coli* strains utilise different compounds,[30,31] which supports the above hypothesis and we have recently demonstrated that UPEC from different STs vary significantly in their metabolic potential, indicating the importance of biochemical variation between isolates.[29]

Secondly, the washing process (even in physiological saline; 0.9% NaCl) may cause variations in the biomass due to stress and lysis. The third and most important area concerns the impact of secretions from the isolates. The exometabolome has been the subject of studies in many different fields, such as clinical medicine[32] and plant-pathology[33] as well as microbiology.[34–36] In functional genomics studies, where the aim is to assign function to genes that lack annotation, metabolic footprinting can be used to assess the effect of gene knockouts on the phenotype of microbial mutants. Thus it has been demonstrated that using analytical techniques such as FT-IR spectroscopy and mass spectrometry (GC-MS and DIMS) it is possible to discriminate between different *E. coli* and yeast isolates with gene mutations.[35,37,38]

All three suggestions could be valid and in order to understand more about why the washing of cells alters the phenotypic information and hence clustering requires further investigation. This could include using less complex media and the supernatant needs to be analysed by higher resolution analysis afforded by mass spectrometry to see if certain metabolites in the bacterial footprint correspond to different ST or antibiotic phenotype.

## Concluding remarks

A modified method of sample preparation for FT-IR spectroscopy was examined and applied to a group of UPEC. The first step was to compare between micro-culturing in a bioscreen and traditional shake flask culture, and equivalent phenotypes were revealed using FT-IR spectroscopy. Following this, further whole organism fingerprinting analyses established that the most information rich and informative FT-IR spectra were generated when biomass was analysed *directly* from the micro-cultures. We therefore have confidence in the validity and reproducibility of this approach and believe it can be utilised to handle large numbers of bacterial isolates without the need for sample washing.

From this study it can be seen that a high throughput micro-culture method coupled with FT-IR spectroscopy

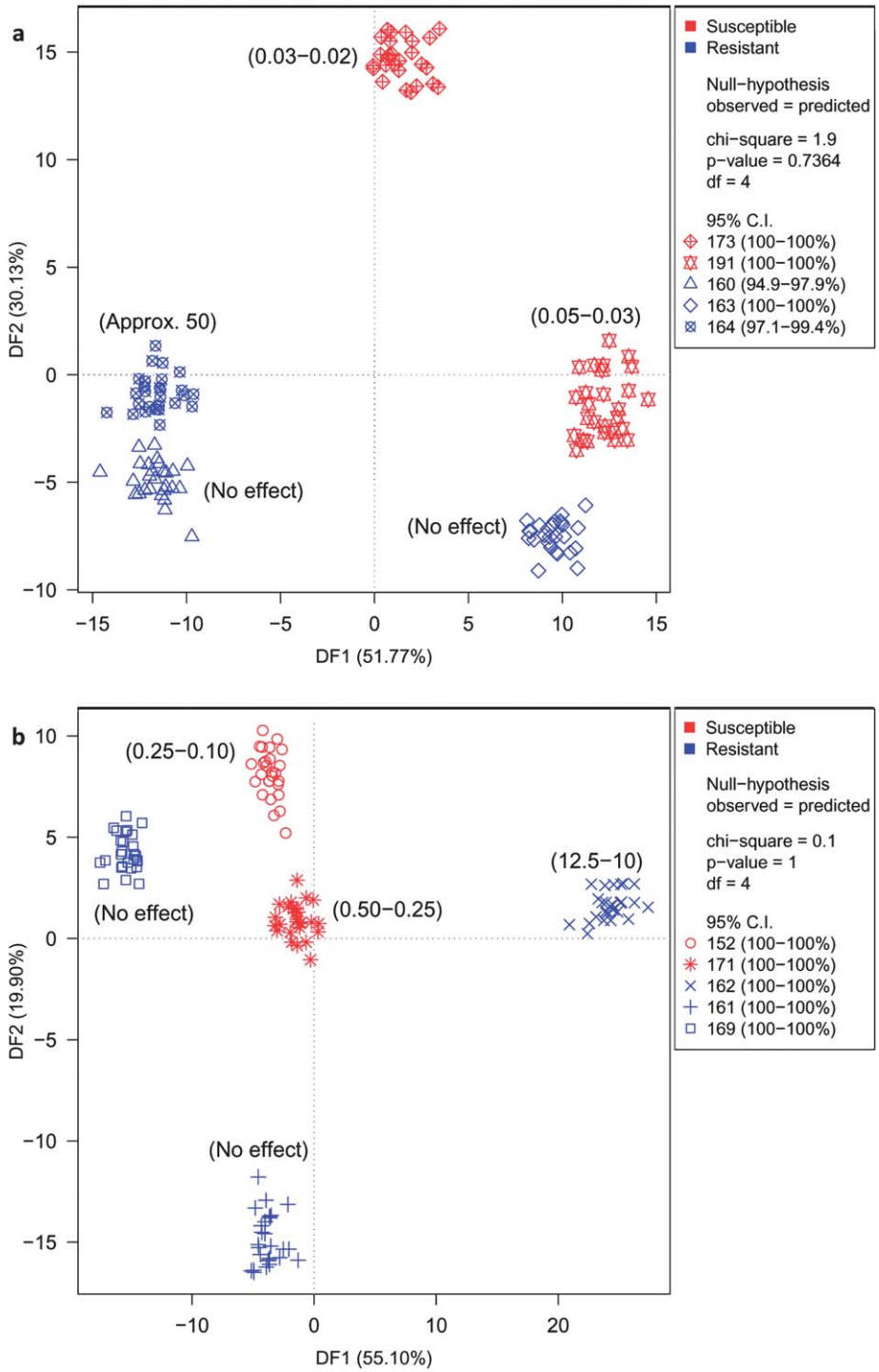


**Fig. 2** PC-DFA scores plot of FT-IR spectra from the non-washed method for (a) ST131 isolates and (b) non-ST131 sensitive isolates. In both cases 30 PCs were extracted from PCA and used as inputs to DFA. Those 30 PCs explain 99.98% of the data variance on both data sets ST131 and non-ST131.

enabled discrimination between uropathogenic bacterial isolates from important clones. Due to the high throughput nature of this approach it has the major advantage that one is now able to analyse several hundred or even thousands of bacterial samples per day. In addition, it is non-destructive and inexpensive, requiring only simple sample preparation. It can therefore be concluded that this method would be suitable for future FT-IR spectroscopic-based investigations of bacteria for identification or characterisation purposes.

HR thanks The Saudi Ministry of higher education and King Saud University for funding. EC and RG are grateful to the EU Commonsense (http://www.fp7projectcommonsense.eu/) project (Grant 261809) financed by the European Commission under the 7th Framework Programme for Research and Technological Development.

## References

1 S. D. Fihn, *N. Engl. J. Med.*, 2003, **349**, 259–266.
2 T. L. Griebling, *J. Urol.*, 2005, **173**, 1281–1287.
3 D. Scholes, T. M. Hooton, P. L. Roberts, A. E. Stapleton, K. Gupta and W. E. Stamm, *J. Infect. Dis.*, 2000, **182**, 1177–1182.
4 A. Baerheim, *Br. Med. J.*, 2001, **323**, 1197–1198.
5 G. Kahlmeter, *J. Antimicrob. Chemother.*, 2003, **51**, 69–76.
6 A. Ronald, *Dis.-Mon.*, 2003, **49**, 71–82.
7 S. H. Lau, S. Reddy, J. Cheesbrough, F. J. Bolton, G. Willshaw, T. Cheasty, A. J. Fox and M. Upton, *J. Clin. Microbiol.*, 2008, **46**, 1076–1080.
8 B. A. Rogers, H. E. Sidjabat and D. L. Paterson, *J. Antimicrob. Chemother.*, 2011, **66**, 1–14.
9 E. K. Kastanos, A. Kyriakides, K. Hadjigeorgiou and C. Pitris, *J. Raman Spectrosc.*, 2009, **41**, 958–963.
10 C. L. Winder and R. Goodacre, *Analyst*, 2004, **129**, 1118–1122.
11 D. Helm, H. Labischinski, G. Schallehn and D. Naumann, *J. Gen. Microbiol.*, 1991, **137**, 69–79.
12 D. Naumann, D. Helm and H. Labischinski, *Nature*, 1991, **351**, 81–82.
13 D. Naumann, in *Encyclopedia of Analytical Chemistry*, ed. R. A. Meyers, John Wiley & Sons, Ltd, Chiceter, 2006, pp. 102–131.
14 B. H. Stuart, *Infrared Spectroscopy: Fundamentals and Applications*, John Wiley & Sons Ltd., Chicester, 2004.
15 S. Garip, F. Bozoglu and F. Severcan, *Appl. Spectrosc.*, 2007, **61**, 186–192.
16 R. Goodacre, E. M. Timmins, P. J. Rooney, J. J. Rowland and D. B. Kell, *FEMS Microbiol. Lett.*, 1996, **140**, 233–239.
17 L. Mariey, J. P. Signolle, C. Amiel and J. Travert, *Vib. Spectrosc.*, 2001, **26**, 151–159.
18 K. Maquelin, C. Kirschner, L. P. Choo-Smith, N. van den Braak, H. P. Endtz, D. Naumann and G. J. Puppels, *J. Microbiol. Methods*, 2002, **51**, 255–271.
19 L. P. Choo-Smith, K. Maquelin, T. van Vreeswijk, H. A. Bruining, G. J. Puppels, N. A. G. Thi, C. Kirschner, D. Naumann, D. Ami, A. M. Villa, F. Orsini, S. M. Doglia, H. Lamfarraj, G. D. Sockalingum, M. Manfait, P. Allouch and H. P. Endtz, *Appl. Environ. Microbiol.*, 2001, **67**, 1461–1469.
20 J. M. Andrews, *J. Antimicrob. Chemother.*, 2001, **48**, 5–16.
21 C. L. Winder, S. V. Gordon, J. Dale, R. G. Hewinson and R. Goodacre, *Microbiology*, 2006, **152**, 2757–2765.
22 H. Martens, J. P. Nielsen and S. B. Engelsen, *Anal. Chem.*, 2003, **75**, 394–404.
23 T. Naes, T. Isaksson and B. Kowalski, *Anal. Chem.*, 1990, **62**, 664–673.
24 R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, Pearson/Prentice Hall, Upper Saddle River, NJ, 2007.
25 B. Manly, *Multivariate Statistical Methods*, Chapman and Hall, London, 1994.
26 W. Windig, J. Haverkamp and P. G. Kistemaker, *Anal. Chem.*, 1983, **55**, 81–88.
27 K. Varmuza and P. Filzmoser, *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, Boca Raton, 2009.
28 B. Efron, *Biometrika*, 1981, **68**, 589–599.
29 T. M. Gibreel, A. R. Dodgson, J. Cheesbrough, F. J. Bolton, A. J. Fox and M. Upton, *J. Clin. Microbiol.*, 2012, **50**, 3202–3207.
30 R. M. M. Klemperer, N. Ismail and M. R. W. Brown, *J. Gen. Microbiol.*, 1979, **115**, 325–331.
31 D. K. Olukoya, *J. Gen. Microbiol.*, 1986, **132**, 3231–3234.
32 W. B. Dunn, M. Brown, S. A. Worton, I. P. Crocker, D. Broadhurst, R. Horgan, L. C. Kenny, P. N. Baker, D. B. Kell and A. E. P. Heazell, *Placenta*, 2009, **30**, 974–980.
33 J. W. Allwood, A. Clarke, R. Goodacre and L. A. J. Mur, *Phytochemistry*, 2010, **71**, 590–597.
34 V. Behrends, T. M. D. Ebbels, H. D. Williams and J. G. Bundy, *Appl. Environ. Microbiol.*, 2009, **75**, 2453–2463.
35 N. N. Kaderbhai, D. I. Broadhurst, D. I. Ellis, R. Goodacre and D. B. Kell, *Comp. Funct. Genomics*, 2003, **4**, 376–391.
36 V. Mapelli, L. Olsson and J. Nielsen, *Trends Biotechnol.*, 2008, **26**, 490–497.
37 J. Allen, H. M. Davey, D. Broadhurst, J. K. Heald, J. J. Rowland, S. G. Oliver and D. B. Kell, *Nat. Biotechnol.*, 2003, **21**, 692–696.
38 S. Mas, S. G. Villas-Boas, M. E. Hansen, M. Akesson and J. Nielsen, *Biotechnol. Bioeng.*, 2007, **96**, 1014–1022.